RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Journal Name

COMMUNICATION

Facile Synthesis of SAM-Peptide Conjugates through Alkyl Linkers Targeting Protein *N*-terminal Methyltransferase 1

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Gang Zhang[a] and Rong Huang*[a]

We report the first chemical synthesis of SAM-peptide conjugates through alkyl linkers to prepare bisubstrate analogs for protein methyltransferases. We demonstrated its application by developing a series of bisubstrate inhibitors for protein N-terminal methyltransferase 1 and the most potent one exhibited a K_i value of 310 ± 55 nM.

Protein methylation has been recognized as an important post-translational modification because of its implications in a variety of disorders including cancer, cardiovascular diseases, inflammation, metabolic and neurodegenerative diseases.[1-5] Consequently, protein methyltransferases that are responsible for this modification have drawn a lot of attention as potential therapeutic targets. The protein methyltransferase family consists of nearly fifty predicted protein lysine methyltransferases (PKMTs), over forty predicted protein arginine methyltransferases (PRMTs), and two newly discovered protein *N*-terminal methyltransferases (NTMTs). They all have two binding pockets: one for a protein substrate and the other for a methyl donor *S*-adenosyl-*L*-methionine (SAM). PKMTs, PRMTs, and NTMTs all utilize SAM as a methyl donor to transfer a methyl group from SAM to the nitrogen atom of the epsilon amine of lysine, guanidino group of arginine, and N-terminal α-amine, respectively. The methyl-transfer reaction generally undergoes an S_N2-like mechanism and involves a ternary complex formation.[6–9]

The number of reports concerning protein methylation has undergone exponential growth since 2000. However, the majority of protein methyltransferases involved in specific physiological reactions remains elusive. Hence, specific probes are important to decipher the biological functions and therapeutic potentials of methyltransferase targets.[10–16] A general approach for the development of specific inhibitors of individual methyltransferases would greatly advance the field.

Bisubstrate analogues that simultaneously occupy the binding sites for both substrates to mimic a transition state and have the potential to offer high selectivity for individual enzymes of large and homologous families such as acetyltransferases, kinases, glycosyltransferases and methyltransferases (**Figure 1**).[17–26] Bisubstrate inhibitors can also be used in cellular studies if linked to cell-penetrating peptides including polyarginine and Tat peptide (GRKKRRQRRRPPQ) to improve the cellular uptake.[23] The protein methyltransferases share a common SAM binding site, therefore it imposes a big challenge for small SAM analogs to discriminate specific target unless there is a unique structural feature like DOT1L.[2] Since 2010, many efforts have been made to develop bisubstrate analogues for PRMTs and PKMTs by covalently linking a SAM analogue with a guanidine functionality through various alkylamino groups or an ethylene linker.[20–22] They exhibited high selectivity and validated the feasibility of using an alkyl group to tether both substrate portions to develop specific inhibitors for protein methyltransferases. However, these prototype bisubstrate inhibitors either did not contain a peptide substrate moiety, or were obtained *in situ* via nitrogen mustard which restricted an ethylene group as the only linker. [27,28] Hence, there is an urgent need to address the synthetic challenge to chemically tether a SAM analog with a specific peptide substrate through various alkyl groups with different length to probe optimal length to generate more potent and specific probe for individual protein methyltransferase. To fulfil this need, we describe herein a new synthetic route to connect a SAM analog with different peptides via alkyl linkers.

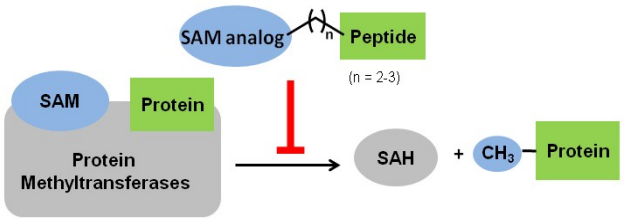


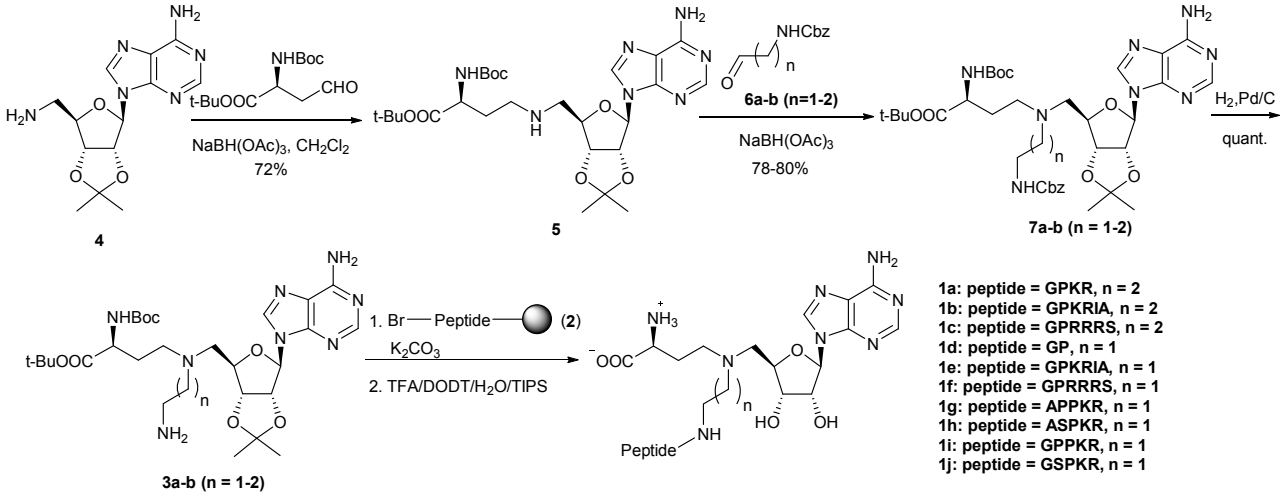
Figure 1. SAM-Peptide conjugates designed to occupy both SAM and protein substrate binding sites

[a] *Department of Medicinal Chemistry, Institute for Structural Biology, Drug Discovery & Development, Virginia Commonwealth University, Richmond, VA, 23219.*

† Electronic Supplementary Information (ESI) available: [organic synthesis, characterization, fluorescence-based assay procedure, and docking study]. See DOI: 10.1039/x0xx00000x

NTMT1, a new addition to the family of protein methyltransferases, recognizes and methylates proteins that have a canonical X-Pro-Lys motif (X = Ala, Gly, Pro, and Ser) at the *N*-terminus.[29,30] So far, there is only one specific inhibitor available for this enzyme, which was synthesized through a click chemistry and reported by our laboratory.[24] To explore a new scaffold, we initiated our efforts to design a series of SAM-peptide conjugates by linking a SAM analogue with peptides that start with either Ala or Gly through an alkyl group as our model system. Our crystal structure suggested that the distance between the sulfonium ion and the α-N-terminal nitrogen atom is 4.7 Å.[30] And for most protein methyltransferases, this

Scheme 1. Synthesis of SAM-Peptide conjugates



distance varies from 2.2 Å to 4.7 Å.[29–33] Hence, we chose an ethylene or a propylene group as a linker and designed a series of compounds **1a-f** targeting NTMT1 with substrate peptides that start with GPK/R.[34, 35] We also designed compounds **1g-j** by adding either an Ala or Gly to an ethylene group to extend the linker length for two peptide substrates PPKR and SPKR to probe an optimal linker. Since the addition of In addition, successful preparation of **1g-j** will further support that our synthetic route may be applied as a general approach to link different peptides to construct bisubstrate analogues for other protein methyltransferases.

Table 1. Inhibition Activities against NTMT1

Compd	n	peptide	IC_{50} (μM)	K_i (μM)
1a	2	GPKR	10 ± 2	2.6 ± 0.7
1b	2	GPKRIA	4.2 ± 1.2	1.1 ± 0.3
1c	2	GPRRRS	0.94 ± 0.16	0.31 ± 0.06
1d	1	GP	77 ± 36	19 ± 9
1e	1	GPKRIA	10 ± 3	2.6 ± 0.8
1f	1	GPRRRS	4.6 ± 1.5	1.2 ± 0.4
1g	1	APPKR	>100	-
1h	1	ASPKR	>100	-
1i	1	GPPKR	>100	-
1j	1	GSPKR	>100	-

We have successfully synthesized these SAM-peptide conjugates (**Scheme 1**). Compound **4** was synthesized from commercially available adenosine after protection of 2'- and 3'-hydroxyl groups by the isopropylidene group[36], conversion to an azide[37], and followed by a hydrogenation reaction.[38] Then **4** reacted with Boc-protected aspartic aldehyde gave **5** by reductive amination.[20] Amines of the amino alcohols were protected with a Cbz group by treatment of benzyl chloroformate (CbzCl).[39] Subsequent oxidation with Dess-Martin periodinane yielded **6a-b**[40], which were subjected to a reductive amination in the presence of **5** and sodium triacetoxyborohydride to produce **7a-b** in 78-80% yield. Removal of the Cbz group by hydrogenation reaction from **7a-b** led to **3a-b**. Compound **3a-b** reacted with α-bromo peptide on resin and was cleaved from the resin with a cleavage cocktail containing trifluoroacetic acid (TFA)/ 2,2'-(Ethylenedioxy)-diethanethiol(DODT)/H_2O/ triisopropylsilane (TIPS) (94:2.5:2.5:1,v/v), and purified by RP-HPLC to afford the target compounds **1a-j**. The α-bromo peptides were prepared by using 2-bromo acetic acid or 2-bromo propionic acid at the *N*-terminal position following general solid-phase peptide synthesis.

We evaluated the inhibitory activities of all synthesized compounds at both K_m values of SAM (8 μM) and RCC1-10 (0.9 μM) in a fluorescence-based assay.[9] Among them, compounds **1a-f** exhibited an IC_{50} value ranging from 780 nM to 113 μM (**Table 1**). In particular, compound **1c** displayed an IC_{50} value of 0.94 ± 0.16 μM and a K_i value of 310 ± 55 nM. Structure-activity relationships indicate that a propylene group is an optimal linker for NTMT1, since **1b-c** are more potent than **1e-f**, respectively. These results further validated our structural discovery of the unique distance of 4.7 Å for NTMT1. As the peptide length extends to six like **1b-c** and **1e-f**, the NTMT1 inhibition activity increases. These results confirmed that the first six amino acids are important for NTMT1 recognition.[9] In addition, peptide sequence specificity contributes to the inhibitory activity. For NTMT1, the second and third positions of peptide substrates are critical for NTMT1 recognition. Normally, Pro is at the second position and Lys or Arg is at the third position,

respectively. As shown in **Table 1**, **1g-j** barely showed any inhibition even at 100 μM, which suggested that addition of Gly or Ala to the linker failed to orient PPKR and SPKR to the peptide binding pocket. The bisubstrate analogues (**1c** and **1f**) with peptide sequence GPRRRS showed the most potent activities than related compounds (**1b** and **1e**), which suggests that GPRRRS can be well recognized by NTMT1. This result correlates with previous study of centromere protein A which contains a GPRRRS peptide at its N-terminus and has been shown to be predominantly trimethylated in mitosis.[34]

In order to assess the selectivity of **1c** to serve as a valuable inhibitor of NTMT1, we evaluated its selectivity *versus* two closely related protein methyltransferases, protein arginine methyltransferase 1 and lysine methyltransferase G9a. At 30 μM, it did not show any significant inhibition of either G9a or PRMT1. We also examined how compound **1c** affects the progression of *N*-α-amine methylation at 5 μM by MALDI-MS.[9,24,41] Triplicate samples of RCC1-10 peptide (SPKRIAKRRS) along with compound **1c** were subjected to NTMT1 methylation assays. Following these assays, samples were analyzed at 20 min to monitor the methylation progression. Dimethylation and trimethylation of RCC1-10 were completely abolished. Monomethylated RCC1-10 was substantially reduced to 19% in the presence of **1c** (**Figure 2**).

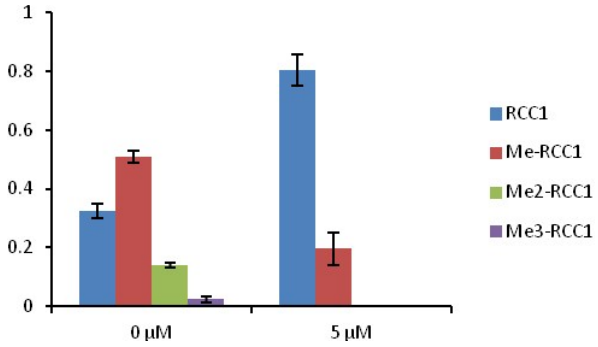


Figure 2. Inhibition on methylation states of RCC1-10 with 5 μM of compound 1c

To understand the interactions between compound **1c** and NTMT1, we performed molecular docking of **1c** into the SAM and peptide substrate binding sites of NTMT1 using Gold 5.2. The result suggested that compound **1c** can occupy both SAM and peptide substrate binding sites simultaneously. The SAM part in **1c** was superimposed well with SAH and retained similar interactions with NTMT1. The Ser6 of the peptide interacts with GLU213 and the guanidino groups on the three Arg residues interact with LEU31, GLY32 and LEU176 (**Figure 3**).

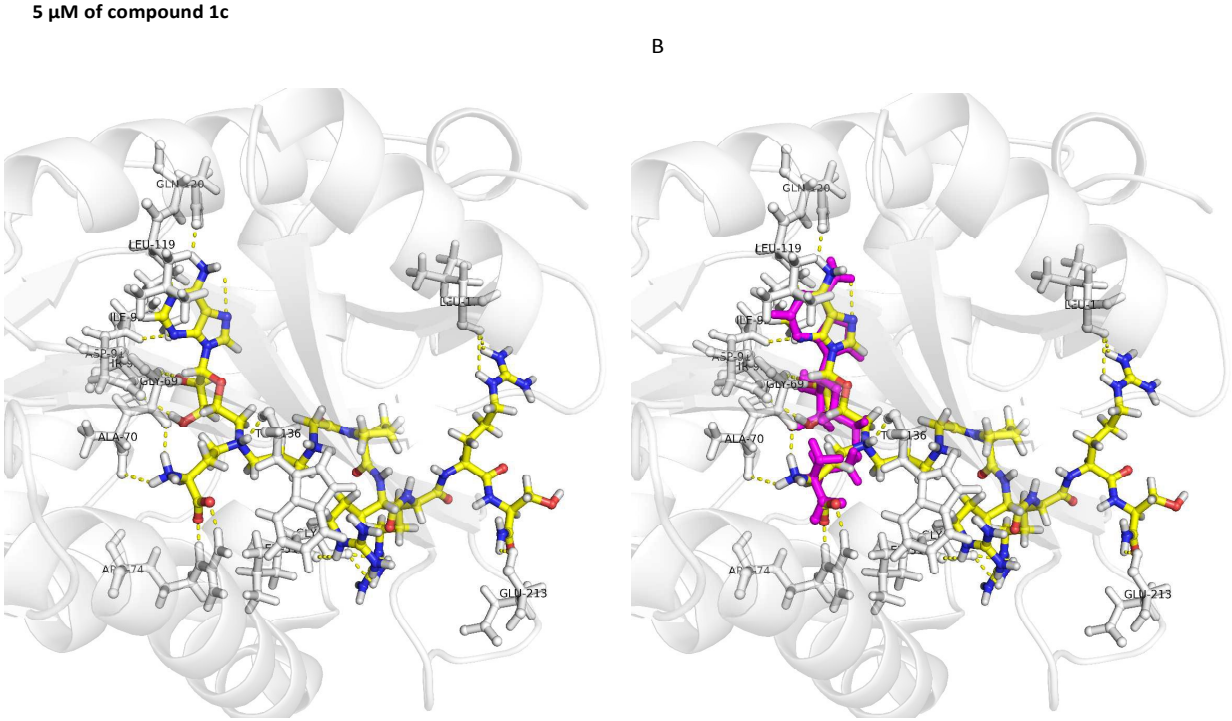


Figure 3. (A) Docking study of 1c with carbons in yellow to crystal structure of NTMT1 complexed with SAH (PDB: 2EX4). (B) Superimposed structure of 1c (carbons in yellow, nitrogens in blue and oxygens in red) with SAH (purple) in the complex. Only interacting residues were labeled. Hydrogen bonds are shown as yellow dotted lines.

In summary, we successfully synthesized a series of SAM-peptide conjugates among which compound **1c** exhibited a K_i value of 235 ± 40 nM for NTMT1. The optimal linker is a propylene group for NTMT1 in this series. This route is the first chemical synthesis to link a SAM analogue with peptide through alkyl linkers. It has the potential to be adapted to build bisubstrate analogues for other PMTs since alkyl linkers have been proven to be viable linkers for both PRMTs and PKMTs.

We thank Dr. Raymond Trievel for the SAH Hydrolase plasmid, and Dr. Yujun Zheng for the PRMT1 plasmid. We appreciate Dr. Darrel L.

Peterson for purifying NTMT1 and G9a. The authors acknowledge financial support from the VCU Presidential Research Quest Fund (R.H.) .

Notes and references

1 M. H. Cho, J. H. Park, H. J. Choi, M. K. Park, H. Y. Won, Y. J. Park, C. H. Lee, S. H. Oh, Y. S. Song, H. S. Kim, Y. H. Oh, J. Y. Lee and G. Kong, *Nat. Commun.*, 2015, **6**, 7821.
2 Y. Yao, P. Chen, J. Diao, G. Cheng, L. Deng, J. L. Anglin, B.V. V. Prasad, and Y. Song, *J. Am. Chem. Soc.*, , 2011, **133**, 16746-16749.
3 X. Liu, L. Wang, H. Li, X. Lu, Y. Hu, X. Yang, C. Huang and D. Gu, *Atherosclerosis*, 2014, **233**, 349–356.
4 W. W. Tsai, S. Niessen, N. Goebel, J. R. Yates, E. Guccione and M. Montminy, *Proc. Natl. Acad. Sci.,* 2013, **110**, 8870–8875.
5 C. Scaramuzzino, I. Casci, S. Parodi, P. M. J. Lievens, M. J. Polanco, C. Milioto, M. Chivet, J. Monaghan, A. Mishra, N. Badders, T. Aggarwal, C. Grunseich, F. Sambataro, M. Basso, F. O. Fackelmayer, J. P. Taylor, U. B. Pandey and M. Pennuto, *Neuron*, 2015, **85**, 88–100.
6 M. F. Hegazi, R. T. Borchard and R. L. Schowen, *J. Am. Chem. Soc.*, 1976, **98**, 3048–3049.
7 M. O'Gara, S. Klimaŝauskas, R. J. Roberts and X. Cheng, *J. Mol. Biol.*, 1996, **261**, 634–645.
8 J. C. Wu and D. V. Santi, *J. Biol. Chem.*, 1987, **262**, 4778–4786.
9 S. L. Richardson, Y. Mao, G. Zhang, P. Hanjra, D. L. Peterson and R. Huang, *J. Biol. Chem.*, 2015, **290**, 11601–11610.
10 R. A. Copeland, E. J. Olhava and M. P. Scott, *Curr. Opin. Chem. Biol.*, 2010, **14**, 505–510.
11 Z. Fan, C. Hao, M. Li, X. Dai, H. Qin, J. Li, H. Xu, X. Wu, L. Zhang, M. Fang, B. Zhou, W. Tian and Y. Xu, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2015, **1849**, 1219–1228.
12 L. L. Ferreira, R. Couto and P. J. Oliveira, *Eur. J. Clin. Invest.*, 2015, **45**, 32–36.
13 D. S. Pan, Q. J. Yang, X. Fu, S. Shan, J. Z. Zhu, K. Zhang, Z. B. Li, Z. Q. Ning and X. P. Lu, *Med. Chem. Commun.*, 2014, **5**, 1789–1796.
14 R. Pereira, R. Benedetti, S. Pérez-Rodríguez, A. Nebbioso, J. García-Rodríguez, V. Carafa, M. Stuhldreier, M. Conte, F. Rodríguez-Barrios, H. G. Stunnenberg, H. Gronemeyer, L. Altucci and Á. R. de Lera, *J. Med. Chem.*, 2012, **55**, 9467–9491.
15 R. A. Copeland, M. E. Solomon and V. M. Richon, *Nat. Rev. Drug Discov.*, 2009, **8**, 724–732.
16 C. Gros, J. Fahy, L. Halby, I. Dufau, A. Erdmann, J.-M. Gregoire, F. Ausseil, S. Vispé and P. B. Arimondo, *Biochimie*, 2012, **94**, 2280–2296.
17 O. D. Lau, T. K. Kundu, R. E. Soccio, S. Ait-Si-Ali, E. M. Khalil, A. Vassilev, A. P. Wolffe, Y. Nakatani, R. G. Roeder and P. A. Cole, *Molecular Cell*, 2000, **5**, 589-595.
18 A. C. Hines and P. A. Cole, *Bioorganic & Medicinal Chemistry Letters*, 2004, **14**, 2951-2954.
19 P. D. Eason and B. Imperiali, *Biochemistry*, 1999, **38**, 5430-5437.
20 J. Dowden, W. Hong, R. V. Parry, R. A. Pike and S. G. Ward, *Bioorganic Med. Chem. Lett.*, 2010, **20**, 2103–2105.
21 J. Dowden, R. A. Pike, R. V Parry, W. Hong, U. A. Muhsen and S. G. Ward, *Org. Biomol. Chem.*, 2011, **9**, 7814.
22 M. van Haren, L. Q. van Ufford, E. E. Moret and N. I. Martin, *Org. Biomol. Chem.*, 2014, **13**, 549–560.
23 B.P. Barnett, Y. Hwang, M. S. Taylor, H. Kirchner, P.T. Pfluger, V. Bernard, Y. Y. Lin, E. M. Bowers, C. Mukherjee, W. J. Song, P. A. Longo, D. J. Leahy, M. A. Hussain, M. H. Tschöp, J. D. Boeke, P. A. Cole. *Science*, 2010, **330**, 1689-92.
24 G. Zhang, S. L. Richardson, Y. Mao and R. Huang, *Org. Biomol. Chem.*, 2015, **13**, 4149–4154.
25 C. D. Chang and J. K. Coward, *J. Med. Chem.*, 1976, **19**, 684–691.
26 M. R. Burns and J. K. Coward, *Bioorg. Med. Chem.*, 1996, **4**, 1455–1470.
27 Y. Du, C. E. Hendrick, K. S. Frye and L. R. Comstock, *Chembiochem*, 2012, **13**, 2225–2233.
28 T. Osborne, R. L. Weller Roska, S. R. Rajski and P. R. Thompson, *J. Am. Chem. Soc.*, 2008, **130**, 4574–4575.
29 C. E. S. Tooley, J. J. Petkowski, T. L. Muratore-Schroeder, J. L. Balsbaugh, J. Shabanowitz, M. Sabat, W. Minor, D. F. Hunt and I. G. Macara, *Nature*, 2010, **466**, 1125–1128.
30 C. Dong, Y. Mao, W. Tempel, S. Qin, L. Li, P. Loppnau, R. Huang and J. Min, *Genes & Dev.*, 2015, **29**, 2343–2348.
31 Y. Yue, Y. Chu and H. Guo, *Molecules*, 2015, **20**, 10032–10046.
32 H. B. Guo and H. Guo, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 8797–8802.
33 X. Zhang, L. Zhou and X. Cheng, *EMBO J.*, 2000, **19**, 3509–3519.
34 A. O. Bailey, T. Panchenko, K. M. Sathyan, J. J. Petkowski, P. J. Pai, D. L. Bai, D. H. Russell, I. G. Macara, J. Shabanowitz, D. F. Hunt, B. E. Black and D. R. Foltz, *Proc. Natl. Acad. Sci.*, 2013, **110**, 11827–11832.
35. Dai, X., Otake, K., You, C., Cai, Q., Wang, Z., Masumoto, H., and Wang, Y, *J. Proteome Res.* 2013, **12**, 4167-4175.
36. A. Hampton, *J. Am. Chem. Soc.*, 1961, **83**, 3640–3645.
37 F. Liu and D. J. Austin, *J. Org. Chem.*, 2001, **66**, 8643–8645.
38 P. Ciuffreda, A. Loseto and E. Santaniello, *Tetrahedron*, 2002, **58**, 5767–5771.
39 C. De Cola, A. Manicardi, R. Corradini, I. Izzo and F. De Riccardis, *Tetrahedron*, 2012, **68**, 499–506.
40 F. Diness, J. Beyer and M. Meldal, *QSAR Comb. Sci.*, 2004, **23**, 130–144.
41. S. L. Richardson, P. Hanjra, G. Zhang, B. D. Mackie, D. L. Peterson, R. Huang. *Anal. Biochem*., 2015, **478**, 59-64.